

EXHIBIT D

EPA/630/P-03/001F
March 2005

Guidelines for Carcinogen Risk Assessment

Risk Assessment Forum
U.S. Environmental Protection Agency
Washington, DC

2.3.5.4. Judging Data

Criteria that are generally applicable for judging the adequacy of mechanistically based data include:

- mechanistic relevance of the data to carcinogenicity,
- number of studies of each endpoint,
- consistency of results in different test systems and different species,
- similar dose-response relationships for tumor and mode of action-related effects,
- conduct of the tests in accordance with generally accepted protocols, and
- degree of consensus and general acceptance among scientists regarding interpretation of the significance and specificity of the tests.

Although important information can be gained from *in vitro* test systems, a higher level of confidence is generally given to data that are derived from *in vivo* systems, particularly those results that show a site concordance with the tumor data.

It is important to remember that when judging and considering the use of any data, the basic standard of quality, as defined by the EPA Information Quality Guidelines, should be satisfied.

2.4. MODE OF ACTION—GENERAL CONSIDERATIONS AND FRAMEWORK FOR ANALYSIS

2.4.1. General Considerations

The interaction between the biology of the organism and the chemical properties of the agent determine whether there is an adverse effect. Thus, mode of action analysis is based on physical, chemical, and biological information that helps to explain key events in an agent's influence on development of tumors. The entire range of information developed in the assessment is reviewed to arrive at a reasoned judgment. An agent may work by more than one mode of action, both at different sites and at the same tumor site. Thus the mode of action and human relevance cannot necessarily be generalized to other toxic endpoints or tissues or cell types without additional analyses (IPCS, 1999; Meek et al., 2003). At least some information

bearing on mode of action (e.g., SAR, screening tests for mutagenicity) is present for most agents undergoing assessment of carcinogenicity, even though certainty about exact molecular mechanisms may be rare.

Information for mode of action analysis generally includes tumor data in humans and animals and among structural analogues, as well as the other key data. The more complete the data package and the generic knowledge about a given mode of action, the more confidence one has and the more one can rely on assessment of available data rather than reverting to default options to address the absence of information on mode of action. Reasoned judgments are generally based on a data-rich source of chemical, chemical class, and tumor type-specific information. Many times there will be conflicting data and gaps in the information base; it is important to carefully evaluate these uncertainties before reaching any conclusion.

In making decisions about potential modes of action and the relevance of animal tumor findings to humans (Ashby et al., 1990; Ashby and Tennant, 1991; Tennant, 1993; IPCS 1999; Sonich-Mullin et al., 2001; Meek et al., 2003), very often the results of chronic animal studies may give important clues. Some of the important factors to review include:

- tumor types, for example, those responsive to endocrine influence or those produced by DNA-reactive carcinogens;
- number of studies and of tumor sites, sexes, and species affected or unaffected in those studies and if the data present a coherent story;
- similarity of metabolic activation and detoxification for a specific chemical between humans and tested species;
- influence of route of exposure on the spectrum of tumors and whether they occur at point of exposure or systemic sites;

- effect of high dose exposures on the target organ or systemic toxicity that may not reflect typical physiological conditions, for example, urinary chemical changes associated with stone formation, effects on immune surveillance;
- presence of proliferative lesions, for example, hepatic foci, or hyperplasia;
- effect of dose and time on the progression of lesions from preneoplastic to benign tumors, then to malignant;
- ratio of malignant to benign tumors as a function of dose and time;
- time of appearance of tumors after commencing exposure;
- development of tumors that invade locally or systemically, or lead to death;
- tumors at organ sites with high or low background historical incidence in laboratory animals;
- biomarkers in tumor cells, both induced and spontaneous, for example, DNA or protein adducts, mutation spectra, chromosome changes, oncogene activation; and/or
- shape of the dose-response curve in the range of tumor observation, for example, linear versus nonlinear.

Some of the myriad ways in which information from chronic animal studies influences mode of action judgments include, but are not limited to, the following:

- multisite and multispecies tumor effects that are often associated with mutagenic agents;

- tumors restricted to one sex or species suggesting an influence restricted to gender, strain, or species;
- late onset of tumors that are primarily benign, are at sites with a high historical background incidence, or show reversal of lesions on cessation of exposure suggesting a growth-promoting mode of action;
- the possibility that an agent acting differently in different tissues; or
- the possibility that has more than one mode of action in a single tissue.

Simple knowledge of sites of tumor increase in rodent studies can give preliminary clues as to mode of action. Experience at the National Toxicology Program (NTP) indicates that substances that are DNA reactive and that produce gene mutations may be unique in producing tumors in certain anatomical sites, whereas tumors at other sites may arise from both mutagenic or nonmutagenic influences (Ashby and Tennant, 1991; Huff et al., 1991).

The types of data and their influence on judgments regarding mode of action are expected to evolve, both as science advances and as the risk assessment community gains more experience with these analyses. This section contains a framework for evaluating hypothesized mode(s) of action. This framework has similarities to and differences with the concepts presented in other MOA frameworks (e.g., IPCS, 1999; Sonich-Mullin et al., 2001; Meek et al., 2003). Differences are often due to the context of the use for the framework. For example, the Meek et al. (2003) presents a stand-alone document for addressing mode of action issues; thus, it recommends that conclusions concerning MOA be rendered separately. In these cancer guidelines, however, they are incorporated into the context of all of the data regarding weight of the evidence for carcinogenicity.

2.4.2. Evaluating an Hypothesized Mode of Action

2.4.2.1. *Peer Review*

In reaching conclusions, the question of "general acceptance" of a mode of action should be tested as part of the independent peer review that EPA obtains for its assessment and conclusions. In some cases the mode of action may already have been established by development of a large body of research information and characterization of the phenomenon over time. In some cases there will have been development of an Agency policy (e.g., mode of action involving alpha-2u-globulin in the male rat [U.S. EPA, 1991b]) or a series of previous assessments in which both the mode of action and its applicability to particular cases has been explored. If so, the assessment and its peer review can focus on the evidence that a particular agent acts in this mode. The peer review should also evaluate the strengths and weaknesses of competing modes of action.

In other cases, the mode of action may not have previously been the subject of an Agency document. If so, the data to support both the mode of action and the associated activity of the agent should undergo EPA assessment and subsequent peer review.

2.4.2.2. *Use of the Framework*

The framework supports a full analysis of mode of action information, but it can also be used as a screen to decide whether sufficient information is available to evaluate or whether the data gaps are too substantial to justify further analysis. Mode of action conclusions are used to address the question of human relevance of animal tumor responses, to address differences in anticipated response among humans, such as between children and adults or men and women; and as the basis of decisions about the anticipated shape of the dose-response relationship. Guidance on the latter appears in Section 3.

This framework is intended to provide an analytical approach for evaluating the mode of action. It is neither a checklist nor a list of required criteria. As the type and amount of information will depend on the mode of action postulated, scientific judgment is important to determine if the weight of evidence is sufficient.

2.4.3. Framework for Evaluating Each Hypothesized Carcinogenic Mode of Action

This framework is intended to be an analytic tool for judging whether available data support a mode of carcinogenic action hypothesized for an agent. It is based upon considerations for causality in epidemiologic investigations originally articulated by Hill (1965) but later modified by others and extended to experimental studies. The original Hill criteria were applied to epidemiologic data, whereas this framework is applied to a much wider assortment of experimental data, so it retains the basic principles of Hill but is much modified in content.

The modified Hill criteria can be useful for organizing thinking about aspects of causation, and they are consistent with the scientific method of developing hypotheses and testing those hypotheses experimentally. During analysis by EPA, and as guidance for peer review, a key question is whether the data to support a mode of action meet the standards generally applied in experimental biology regarding inference of causation.

All pertinent studies are reviewed in analyzing a mode of action, and an overall weighing of evidence is performed, laying out the strengths, weaknesses, and uncertainties of the case as well as potential alternative positions and rationales. Identifying data gaps and research needs is also part of the assessment.

To evaluate whether an hypothesized mode of action is operative, an analysis starts with an outline of the scientific findings regarding the hypothesized key events leading to cancer, and then weighing information to determine whether there is a causal relationship between these events and cancer formation, i.e., that the effects are critical for induction of tumors. It is not generally expected that the complete sequence will be known at the molecular level. Instead, empirical observations made at different levels of biological organization—biochemical, cellular, physiological, tissue, organ, and system—are analyzed.

Several important points should be considered when working with the framework:

- The topics listed for analysis should *not* be regarded as a checklist of necessary “proofs.” The judgment of whether an hypothesized mode of action is supported by available data takes account of the analysis as a whole.

- The framework provides a structure for organizing the facts upon which conclusions as to mode of action rest. The purpose of using the framework is to make analysis transparent and to allow the reader to understand the facts and reasoning behind a conclusion.
- The framework does not dictate an answer. The weight of evidence that is sufficient to support a decision about a mode of action may be less or more, depending on the purpose of the analysis, for example, screening, research needs identification, or full risk assessment. To make the reasoning transparent, the purpose of the analysis should be made apparent to the reader.
- Toxicokinetic studies may contribute to mode of action analysis by contributing to identifying the active form(s) of an agent that is central to the mode of action. Apart from contributing in this way, toxicokinetics studies may reveal effects of saturation of metabolic processes. These may not be considered key events in a mode of action, but they are given separate consideration in assessing dose metrics and potential nonlinearity of the dose-response relationship.
- Generally, “sufficient” support is a matter of scientific judgment in the context of the requirements of the decisionmaker or in the context of science policy guidance regarding a certain mode of action.
- Even when an hypothesized mode of action is supported for a described response in a specific tissue, it may not explain other tumor responses observed, which should get separate consideration in hazard and dose-response assessment.

For each tumor site being evaluated, the mode of action analysis should begin with a description of the relevant data and key events that may be associated with an hypothesized mode of action and its sequence of key events (see Section 2.4.3.1). This can be followed by a

discussion of various aspects of the experimental support for hypothesized mode(s) of action in animals and humans (see Section 2.4.3.2). The possibility of other modes of action also should be considered and discussed (see Section 2.4.3.3); if there is evidence for more than one mode of action, each should receive a separate analysis. Conclusions about each hypothesized mode of action should address whether the mode of action is supported in animals and is relevant to humans and which populations or lifestages can be particularly susceptible (see Section 2.4.3.4). In a risk assessment document, the analysis of an hypothesized mode of action can be presented before or with the characterization of an agent's potential hazard to humans.

2.4.3.1. Description of the Hypothesized Mode of Action

Summary description of the hypothesized mode of action. For each tumor site, the mode of action analysis begins with a description of the hypothesized mode of action and its sequence of key events. If there is evidence for more than one mode of action, each receives a separate analysis.

Identification of key events. In order to judge how well data support involvement of a key event in carcinogenic processes, the experimental definition of the event or events should be clear and reproducible. To support an association, experiments should define and measure an event consistently.

- Can a list of events be identified that are key to the carcinogenic process?
- Are the events well defined?

Pertinent observations may include, but are not limited to, receptor-ligand changes, cytotoxicity, cell cycle effects, increased cell growth, organ weight differences, histological changes, hormone or other protein perturbations, or DNA and chromosome effects.

2.4.3.2. Discussion of the Experimental Support for the Hypothesized Mode of Action

The experimental support for the hypothesized mode of action should be discussed from several viewpoints patterned after the Hill criteria (see Section 2.2.1.7). For illustration, the explanation of each topic includes typical questions to be addressed to the available empirical data and experimental observations anticipated to be pertinent. The latter will vary from case to case. For a particular mode of action, certain observations may be established as essential in practice or policy, for example, measures of thyroid hormone levels in supporting thyroid hormone elevation as a key event in carcinogenesis.

Strength, consistency, specificity of association. A statistically significant association between events and a tumor response observed in well-conducted studies is generally supportive of causation. Consistent observations in a number of such studies with differing experimental designs increase that support, because different designs may reduce unknown biases. Studies showing "recovery," i.e., absence or reduction of carcinogenicity when the event is blocked or diminished, are particularly useful tests of the association. Specificity of the association, without evidence of other modes of action, strengthens a causal conclusion. A lack of strength, consistency, and specificity of association weakens the causal conclusions for a particular mode of action.

- What is the level of statistical and biological significance for each event and for cancer?
- Do independent studies and different experimental hypothesis-testing approaches produce the same associations?
- Does the agent produce effects other than those hypothesized?
- Is the key event associated with precursor lesions?

Pertinent observations include tumor response associated with events (site of action logically relates to event[s]), precursor lesions associated with events, initiation-promotion studies, and stop/recovery studies.

Dose-response concordance. If a key event and tumor endpoints increase with dose such that the key events forecast the appearance of tumors at a later time or higher dose, a causal association can be strengthened. Dose-response associations of the key event with other precursor events can add further strength. Difficulty arises when an event is not causal but accompanies the process generally. For example, if tumors and the hypothesized precursor both increase with dose, the two responses will be correlated regardless of whether a causal relationship exists. This is similar to the issue of confounding in epidemiologic studies. Dose-response studies coupled with mechanistic studies can assist in clarifying these relationships.

- What are the correlations among doses producing events and cancer?

Pertinent observations include, but are not limited to, 2-year bioassay observation of lesions correlated with observations of hormone changes and the same lesions in shorter term studies or in interim sacrifice.

Temporal relationship. If an event is shown to be causally linked to tumorigenesis, it will precede tumor appearance. An event may also be observed contemporaneously or after tumor appearance; these observations may add to the strength of association but not to the temporal association.

- What is the ordering of events that underlie the carcinogenic process?
- Is this ordering consistent among independent studies?

Pertinent observations include studies of varying duration observing the temporal sequence of events and development of tumors.

Biological plausibility and coherence. It is important that the hypothesized mode of action and the events that are part of it be based on contemporaneous understanding of the biology of cancer to be accepted. If the body of information under scrutiny is consistent with other examples (including structurally related agents) for which the hypothesized mode of action is accepted, the case is strengthened. Because some modes of action can be anticipated to evoke effects other than cancer, the available toxicity database on noncancer effects, for example, reproductive effects of certain hormonal disturbances, can contribute to this evaluation.

- Is the mode of action consistent with what is known about carcinogenesis in general and for the case specifically?
- Are carcinogenic effects and events consistent across structural analogues?
- Is the database on the agent internally consistent in supporting the purported mode of action, including relevant noncancer toxicities?

Pertinent observations include the scientific basis for considering an hypothesized mode of action generally, given the contemporaneous state of knowledge of carcinogenic processes; previous examples of data sets showing the mode of action; data sets on analogues; and coherence of data in this case from cancer and noncancer toxicity studies.

2.4.3.3. Consideration of the Possibility of Other Modes of Action

The possible involvement of more than one mode of action at the tumor site should be considered. Pertinent observations that are not consistent with the hypothesized mode of action can suggest the possibility of other modes of action. Some pertinent observations can be consistent with more than one mode of action. Furthermore, different modes of action can operate in different dose ranges; for example, an agent can act predominantly through cytotoxicity at high doses and through mutagenicity at lower doses where cytotoxicity may not occur.

If there is evidence for more than one mode of action, each should receive a separate analysis. There may be an uneven level of experimental support for the different modes of action. Sometimes this can reflect disproportionate resources spent on investigating one particular mode of action and not the validity or relative importance of the other possible modes of action. Ultimately, however, the information on all of the modes of action should be integrated to better understand how and when each mode acts, and which mode(s) may be of interest for exposure levels relevant to human exposures of interest.

2.4.3.4. Conclusions About the Hypothesized Mode of Action

Conclusions about the hypothesized mode of action should address the issues listed below. For those agents for which the mode of action is considered useful for the risk assessment, the weight of the evidence concerning mode of action in animals as well as its relevance for humans would be incorporated into the weight of evidence narrative (Section 2.5).

(a) Is the hypothesized mode of action sufficiently supported in the test animals?

Associations observed between key events and tumors may or may not support an inference of causation. The conclusion that the agent causes one or more key events that results in tumors is strengthened as more aspects of causation are satisfied and weakened as fewer are satisfied. Consistent results in different experiments that test the hypothesized mode of action build support for that mode of action. Replicating results in a similar experiment does not generally meaningfully strengthen the original evidence, and discordant results generally weaken that support. Experimental challenge to the hypothesized mode of action, where interrupting the sequence of key events suppresses the tumor response or enhancement of key events increases the tumor response, creates very strong support for the mode of action.

(b) Is the hypothesized mode of action relevant to humans? If an hypothesized mode of action is sufficiently supported in the test animals, the sequence of key precursor events should be reviewed to identify critical similarities and differences between the test animals and humans. The question of concordance can be complicated by cross-species differences in toxicokinetics or

toxicodynamics. For example, the active agent can be formed through different metabolic pathways in animals and humans. Any information suggesting quantitative differences between animals and humans is flagged for consideration in the dose-response assessment. This includes the potential for different internal doses of the active agent or for differential occurrence of a key precursor event.

“Relevance” of a potential mode of action is considered in the context of characterization of hazard, not level of risk. Anticipated levels of human exposure are not used to determine whether the hypothesized mode of action is relevant to humans. Exposure information is integrated into the overall risk characterization.

The question of relevance considers all populations and lifestages. It is possible that the conditions under which a mode of action operates exist primarily in a particular population or lifestage, for example, in those with a pre-existing hormonal imbalance. Other populations or lifestages may not be analogous to the test animals, in which case the question of relevance would be decided by inference.

Special attention should be paid to whether tumors can arise from childhood exposure, considering various aspects of development during these lifestages. Because the studies that support a mode of action are typically conducted in mature animals, conclusions about relevance during childhood generally rely on inference. There is currently no standard Agency position regarding the issue of whether tumors arising through the hypothesized mode of action are relevant during childhood; understanding the mode of action implies that there are sufficient data (on either the specific agent or the general mode of action) to form a confident conclusion about relevance during childhood.

(c) Which populations or lifestages can be particularly susceptible to the hypothesized mode of action? If an hypothesized mode of action is judged relevant to humans, information about the key precursor event(s) is reviewed to identify populations or lifestages that might reasonably expected to be particularly susceptible to their occurrence. Although agent-specific data would provide the strongest indication of susceptibility, this review may also rely on general knowledge about the precursor events and characteristics of individuals susceptible to these

events. Any information suggesting quantitative differences between populations or lifestages should be flagged for consideration in the dose-response assessment (see Section 3.5). This includes the potential for a higher internal dose of the active agent or for an increased occurrence of a key precursor event. Quantitative differences may result in separate risk estimates for susceptible populations or lifestages.

The possibility that childhood is a susceptible period for exposure should be explicitly addressed. Generic understanding of the mode of action can be used to gauge childhood susceptibility, and this determination can be refined through analysis of agent-specific data.

2.4.4 *Evolution with Experience*

Several groups have proposed or incorporated mode of action into their risk assessments (see, e.g., U.S. EPA, 1991b; Sonich-Mullin et al., 2001; Meck et al., 2003). As the frameworks and mandates under which these evaluations were produced differ, the specific procedures described in and conclusions drawn may also differ. Nevertheless, the number of case studies from all venues remains limited. More experience with differing modes of action are expected to highlight and illustrate the strengths and limitations of the general framework proposed in these cancer guidelines. Moreover, additional toxicological techniques may expand or change scientific judgments regarding which information is useful for mode of action determinations. As warranted, additional guidance may be proposed as experience is gained and/or as toxicological knowledge advances.

2.5. WEIGHT OF EVIDENCE NARRATIVE

The *weight of evidence narrative* is a short summary (one to two pages) that explains an agent's human carcinogenic potential and the conditions that characterize its expression. It should be sufficiently complete to be able to stand alone, highlighting the key issues and decisions that were the basis for the evaluation of the agent's potential hazard. It should be sufficiently clear and transparent to be useful to risk managers and non-expert readers. It may be useful to summarize all of the significant components and conclusions in the first paragraph of the narrative and to explain complex issues in more depth in the rest of the narrative.